

· 康复工程 ·

基于语音与肌电融合的多自由度假肢控制系统研究

邹国栋 熊平 武振兴 田岚 耿艳娟 李光林

目前,许多国家都在利用各种先进方法研究开发高性能多功能假肢。除装饰假肢和索控假肢之外,从信号源方面分类,现有的假肢大致可以分为脑机接口假肢、周围神经接口假肢、肌电图(electromyography, EMG)及其它生物信号控制假肢几大类。大多数商业化假肢都是利用肢体残留肌肉产生的电信号进行控制,操控假肢时,操控者不仅要记住动作的切换方法及切换顺序,还需要在操控过程中知道当前的动作模式,这无疑给假肢使用者增加了控制难度。由于部分截肢者肢体残留肌肉有限或完全丧失,不能为多功能假肢控制提供足够的有效肌电信息,难以实现多自由度假肢的灵活控制。因此,扩充假肢控制信号源成为假肢控制亟需解决的问题。语音识别技术是一项比较成熟的技术,它能够提供更多较少的信息,可扩展性强,现有的语音集成模块还具有高集成性等特点,这些都使得语音识别技术被广泛运用在残疾人康复领域,如手术机器人、轮椅等^[1]。本研究将截肢者的语音信息作为假肢的控制信息之一,与截肢者残留肌肉产生的肌电信息融合,实现多自由度假肢的控制,旨在验证语音与肌电融合假肢控制系统的有效性。

对象与方法

一、实验对象

征集 7 例肢体健全者和 3 例截肢者参与实验。肢体健全者中,男 4 例,女 3 例,平均年龄(25.4 ± 2.1)岁,平均体重(59.4 ± 9.7)kg,平均身高(168.2 ± 7.2)cm,均为右利手;3 例截肢受试者均为男性,1 例右臂前臂截肢,2 例左臂前臂截肢,平均年龄(34.3 ± 8.1)岁,平均体重(65.0 ± 8.2)kg,平均身高(171.3 ± 1.2)cm。

二、研究方法

1. 语音与肌电信息融合的假肢控制方案:本研究通过语音识别实现假肢的关节切换,用一对残留肌肉的电信号实现已选定关节内动作的类型选择及动作执行(图 1)。系统能完成手张开、手闭合、腕内收、腕外展、屈肘、伸肘六个动作的操控。

本研究所涉及的语音识别仅需要辨别几个简单的词汇。训练时,使用者只需要记住并发出代表机械手臂(三自由度假肢)自由度关键词(“肘”、“腕”、“手”)的指令。系统处理完采集到的语音信号后,训练出一个语音分类器。操控假肢时,当使用者再次发出指令,语音分类器就能够识别出他想要控制的自由度,实现假肢关节的切换。语音信号在本系统中起到协同控制的作用,它只需要控制假肢关节的切换,对于在关节内不

同动作模式的切换及动作的起止,则由肌电信号控制。例如,想要完成手张开这个动作,需要假肢使用者给系统发出“手”的指令,然后控制自己相应的肌肉收缩,从而完成这个动作。其控制示意图见图 1。

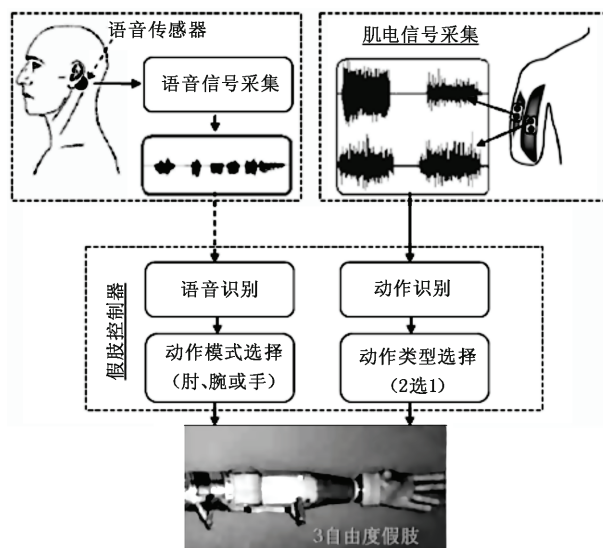


图1 多自由度假肢语音与肌电信号融合控制方案示意图

2. 有声语音识别系统:本研究中,有声语音采集系统的声电转换采用小巧的全指向型背极式驻极体传声器,灵敏度为(-44 ± 3)dB, ($V_s = 3.0$ V, $R_L = 2.2$ k Ω),频率范围为 20 ~ 16 kHz。语音信号经二级放大、二阶巴特沃兹带通滤波、模数转换后,输出数字信号给电脑(图 2)。

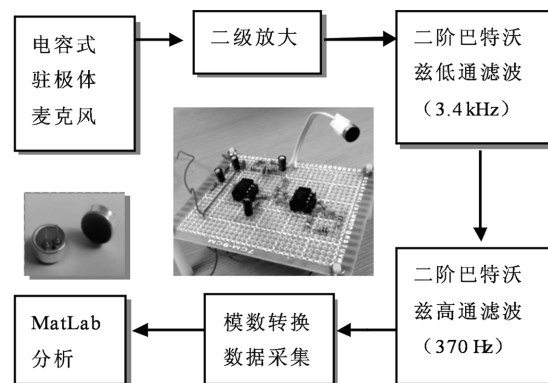


图2 语音信号采集系统框架示意图

本系统涉及的语音识别算法主要是为实现对某一自由度关键词的辨识,如“手”、“腕”、“肘”等。因此,语音识别系统采用对孤立单词识别率较高的动态时间归整(dynamic time warping, DTW)算法。首先将采集到的孤立单词语音信号数据通过计算短时能量和短时过零率的方法截取出有效的语音信号。

DOI:10. 3760/cma. j. issn. 0254-1424. 2012. 010. 022

基金项目:深圳市基础研究重点项目资助(JC201005270295A)

作者单位:410083 长沙,中南大学地球科学与信息物理学院(邹国栋、熊平);中国科学院深圳先进技术研究院(武振兴、田岚、耿艳娟、李光林)

通信作者:熊平,Email:xiongping1025@163.com

数据经加窗分帧后,通过离散傅里叶变换(discrete Fourier transformation, DFT),求出能量谱,通过麦尔频率滤波器组,计算出对数能量,将每个滤波器的输出取对数得出功率谱,并进行反离散余弦变换,得到麦尔频率倒谱系数(Mel frequency cepstral coefficient, MFCC)特征参数,组建成参考模板^[2]。系统将针对所有关键词求出一组麦尔频率倒谱系数特征参数模板。当操作者再次发出语音指令时,系统对测试数据处理后通过 DTW 算法与该组参考模板中的每一个元素进行匹配,选择出匹配度最高的元素作为识别结果输出^[3]。

3. 肌电信号分析系统:肌电信号采集系统由电极、前置放大器及主放大器、数据采集卡、解码分析程序等几部分构成。肌电信号经过调理、模数转换等处理后(图 3),可以解码得出运动起止时间、运动类型以及肌肉紧张程度等信息。

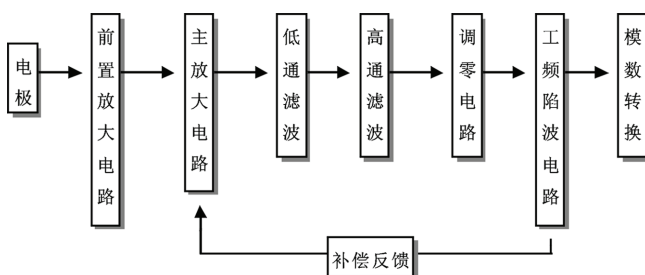


图 3 肌电信号采集框架图

将肌电信号提取出来后,计算出单位分析窗内肌电信号的平均绝对值、过零率、波形长度和正/负斜率变化率四个特征值^[4],通过建立在最小错误率贝叶斯决策规则基础上的 EMG 信号线性判别分析法(linear discriminant analysis, LDA)^[5],对每一种运动类型的 EMG 数据,分别求出该类型所包含的每个分析窗口对应于所有运动类型的后验概率,求得并比较所有运动类型后验概率的值,将后验概率最大的一种运动类型作为该分析窗口的归属运动类型^[6]。在测试过程中,将从测试数据中得到的每一个特征向量通过 LDA 算法计算,选取后验概率值最大的运动类型作为该部分数据的判别运动类型,如果判别运动类型与该部分数据所属的意图运动类型一致,则说明此次分类结果正确,分类器能够正确的识别出该部分信号的运动类型,控制假肢做出正确的动作;反之,则说明此次分类结果错误,假肢无法按照使用者的意图做出希望的动作。

4. 软件系统:本研究自主研发出一套基于矩阵实验室(MatLab)的软件系统(图4)集语音采集与识别、肌电采集与识

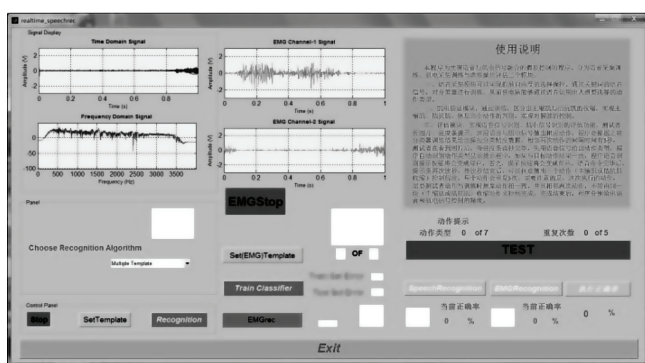


图 4 语音与肌电信息融合的假肢操作系统界面

别、实时控制和操控性能评估于一体,满足本研究提出的各种需求。受试者通过本软件的语音信号模板训练、肌电信号模板训练之后,就能够灵活操控假肢。图形交互界面的使用让假肢的操控及其训练变得简单化、人性化。

5. 实验设计:实验系统的信号采集分为语音信号采集和肌电信号采集两个部分。语音信号采集受试者自身语音信号;对于肢体健全者,肌电信号的采集模拟左前臂截肢的情况,将 2 个表面肌电电极分别放置于受试者左手前臂桡侧腕屈肌(图 5a)和指总伸肌(图 5b)肌腹处,距肘关节 7 cm 左右。对于截肢者,则将 2 个表面肌电电极分别放置于受试者截肢后残留前臂的桡侧腕屈肌(图 5c)和指总伸肌(图 5b)肌腹处。

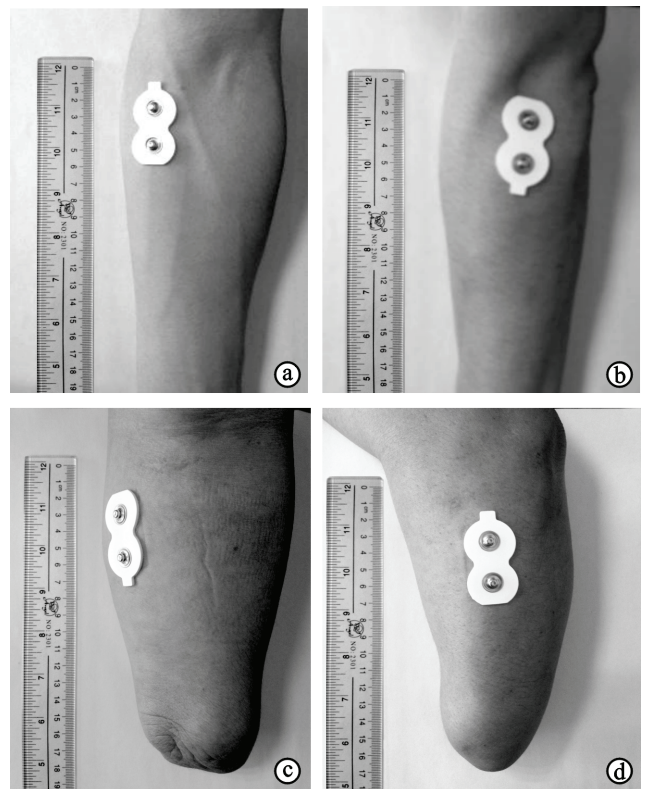


图 5 肌电电极贴放位置示意图

实验时,让受试者分别训练好语音分类器和肌电分类器后,通过重复 5 组“手张开”、“握拳”、“腕内旋”、“腕外旋”、“屈肘”、“伸肘”和“休息”七个动作来检测语音信息与肌电信息融合的假肢控制系统的操控性能。实验要求每位受试者分别进行 3 次训练和测试,并对实验结果取平均值。

三、评估标准及统计

本研究主要通过语音信号识别率、肌电信号识别率和动作完成率三个重要指标来评估多功能假肢的操控性能。

语音信号识别率:语音信号正确识别的比例。它是假肢控制的精度指标之一,它主要体现系统产生正确运动模式分类信息的能力。

肌电信号识别率:即肢体动作的识别率。系统训练完成后,通过肌电信号对受试者关节内动作状态进行判断,当运动模式与目标动作一致时,系统判定受试者动作正确。它主要体现系统对运动状态的操控能力。

动作完成率:在规定时间内,正确完成 5 组“手张开”、“握拳”、“腕内旋”、“腕外旋”、“屈肘”、“伸肘”和“休息”七个动作的百分比。如果在规定时间内,受试者正确完成了所要求的全部动作,则动作完成率为 100%。执行动作时,要求语音信号和肌电信号两项指标都识别正确,才能完成目标动作。此项指标主要衡量系统对假肢的综合操控性能。

对每位受试者进行 3 次独立的控制实验,每次控制实验均需受试者重新训练语音分类器和肌电分类器,每次试验让受试者都完成 5 组“手张开”、“握拳”、“腕内旋”、“腕外旋”、“屈肘”、“伸肘”和“休息”七个动作,统计其语音信号识别率、肌电信号识别率和动作完成率,并对 3 次试验结果数据求均值。

结 果

统计学结果显示,7 例肢体健全受试者的语音信号识别率平均值为 $(92.56 \pm 6.15)\%$,肌电信号识别率平均值为 $(96.87 \pm 1.10)\%$,动作完成率平均值为 $(91.28 \pm 5.25)\%$,部分受试者的动作完成率达 95% 以上。详见表 1。

表 1 7 例肢体健全受试者多功能假肢操控各项性能指标的平均值(%)

受试者	性别	语音信号识别率	肌电信号识别率	动作完成率
受试者 A	男	97.90	97.14	95.23
受试者 B	男	87.78	97.14	87.61
受试者 C	男	81.11	95.25	81.90
受试者 D	女	98.89	98.09	97.14
受试者 E	女	91.11	95.23	87.61
受试者 F	男	98.89	97.14	96.19
受试者 G	女	92.22	98.09	93.28

3 例截肢受试者(表 2)的语音信号识别率平均值为 $(94.81 \pm 3.43)\%$,肌电信号识别率平均值为 $(86.61 \pm 8.90)\%$,动作完成率平均值为 $(84.12 \pm 7.47)\%$ 。详见表 2。

表 2 3 例截肢受试者多功能假肢操控各项性能指标的平均值(%)

受试者	性别	语音信号识别率	肌电信号识别率	动作完成率
截肢者 A	男	90.00	91.42	85.71
截肢者 B	男	97.77	94.28	92.38
截肢者 C	男	96.66	74.14	74.28

考虑到语音信号受环境因素的干扰较大,为了避免假肢的使用者因为他人的语音命令而出现误操作,本研究在处理语音数据时,将关键词分为词汇表内词汇(假肢操作者的命令词汇)和词汇表外词汇(非关键词汇与他人语音词汇),引入了支持向量机(support vector machine, SVM)分类器^[8]与 DTW 算法结合,用于提高词汇表内关键词的识别率和词汇表外词汇的拒识率。对此,本研究另选取 8 例受试者,选取 15 个关键词作为词汇表内词汇,37 个与关键词读音相近的汉语单词及环境噪声片段作为词汇表外词汇,每个单词重复 10 次,组建了一个 4160 个样本的集合(包含训练样本与测试样本),并对此进行离线数据分析。结果如表 3 所示。

表 3 语音数据进行 SVM 预测后语音信号(离线分析)的拒识率和识别率(%)

受试者	S1	S2	S3	S4	S5	S6	S7	S8	识别率
S1	95.7	95.8	77.9	74.9	43.6	95.8	94.9	74.3	91.3
S2	98.2	97.8	85.7	83.9	46.9	90.7	94.0	87.2	90.0
S3	97.0	98.8	97.3	81.2	48.7	85.7	96.4	80.9	97.3
S4	97.0	98.5	83.0	93.0	37.3	89.3	97.9	83.6	80.7
S5	96.4	98.8	84.2	88.7	96.8	93.1	99.1	72.8	97.3
S6	96.7	99.4	73.1	78.8	45.1	93.0	94.6	90.4	88.0
S7	96.1	94.0	89.6	79.7	49.3	85.4	93.5	91.3	90.0
S8	99.1	88.1	85.1	81.8	45.4	85.4	98.2	99.5	100.0
平均值	97.2	96.2	82.7	81.3	45.2	89.3	96.4	82.9	91.8

注:阴影部分表示系统对同一受试者语音信息的词汇表外词汇拒识率,其它数据均为受试者模板对非己语音信息的拒识率;识别率为系统对受试者词汇表内关键词的识别率;拒识率平均值均为受试者模板对非己语音信息的平均拒识率

讨 论

由假肢控制实验结果可以看出,健全受试者和截肢受试者的语音信号平均识别率高达 $(93.23 \pm 5.57)\%$,语音信号的高识别率验证了它作为假肢操控系统扩展信号源的可行性;由于肢体健全受试者和截肢者均无语言障碍,健全受试者和截肢受试者语音信号识别率的差别并不大。此项指标真实地反映了系统对语音关键词的识别率。

肌电信号方面,本研究肢体健全受试者平均识别率为 $(96.87 \pm 1.10)\%$,截肢者平均识别率为 $(86.61 \pm 8.90)\%$ 。造成这一差异的原因是截肢者截肢后,残留肌肉长期得不到锻炼,肌肉严重萎缩,肌电信号非常微弱,而 2 个电极需要还原出 3 个动作状态,这种情况势必要求肌电信号分类器有很高的分类精度。肌电信号识别率的差异说明由于个体的差异,传统单一的肌电信号模式识别在假肢操控方面会存在很大的局限性;肌电信号微弱的截肢者往往很难控制好单一信号源的肌电假肢。

动作完成的正确需要语音信号和肌电信号均识别正确,所以动作完成率是最终体现假肢稳定程度和操控性能的指标。健全受试者动作完成率平均值为 $(91.28 \pm 5.25)\%$,截肢受试者动作完成率平均值为 $(84.12 \pm 7.47)\%$,在语音信号源无差别的情况下,这种差异主要是受残疾人的肌电信号识别率影响造成的。

总之,本研究整个操控方案无论对于健全受试者还是对于截肢受试者,都有不错的识别率,整体识别率保持在 70% 以上。需要指出的是,本研究中,截肢受试者 C,截肢年限长,肌肉萎缩很严重,根本无法使用常规假肢,但是在此操控方案中,动作完成率亦能达到 74.28%。可见,语音信息和肌电信息融合的假肢控制系统具有良好的操控性能。

在拒识率方面,从本研究表 3 中可以看出,加入了支持向量机对语音数据进行预测分析后,语音信号的平均识别率达到 $(91.8 \pm 5.8)\%$,同一受试者语音信息的词汇表外词汇拒识率为 $(95.8 \pm 2.2)\%$,受试者模板对非己语音信息的拒识率为 $(83.9 \pm 15.9)\%$ 。SVM 分类器与 DTW 算法结合对词汇表外的语音数据和对非模板训练者数据的离线拒识率很高,在未来

的研究中,可以将此算法加入到假肢控制系统中,提高假肢控制的抗干扰能力。

然而,以上结果均在实验室环境中得出,在生活环境往往存在诸多的干扰因素,如何提高假肢操控系统的抗干扰能力、避免假肢误操作是本研究下一步的工作重点。

在下一步的研究中,需从以下几个方面对此系统进行改进:①采用空气作为传播媒质的有声语音总会受到日常生活中环境噪音的干扰,故可考虑利用假肢使用者的无声耳语(non-audible murmur)取代有声语音操控假肢^[9]。无声耳语是一种无声带振动的信号,它包含的信息可通过人体软组织的振动进行传导,并可以由佩戴于隐蔽部位(如耳后乳突部位)的传感器捕捉到。由于不使用空气作为传播媒质且不需要使用者真正发出声音,故此方法不仅可以去除周围嘈杂环境对语音信号的干扰,还能避免假肢使用者对着假肢“说话”的尴尬。②截肢者的残肢长期得不到锻炼,肌肉萎缩严重,肌电信号极其微弱。针对此类截肢者可用肌肉数量的不同,尝试用更少的肌电信号(1个)和语音信号融合来控制假肢。③在系统集成方面,可以将语音传感器佩戴于假肢使用者的头部,并使用无线技术和控制器通讯,摆脱有线的束缚;将语音信号识别模块、肌电信号识别模块及假肢驱动模块集成于控制器中,与表面肌电电极一同置于假肢接受腔内,方便假肢使用者的佩戴及操作。

参 考 文 献

- [1] Obando MA, Payne JH. The future application of the robotic arm (automatic endoscopic system for optimal positioning or AESOP) with voice recognition in sinus endoscopic surgery. Oper Techn Otolaryngol-Head and Neck Surg, 2003, 14: 55-57.
- [2] 苏昊, 王民, 李宝. 一种改进的 DTW 语音识别系统. 中国西部科技, 2011, 10: 38-39.
- [3] 张钢, 朱铮涛, 何淑贤. 应用 DTW 的语音(声纹)鉴别技术研究. 中国测试技术, 2007, 33: 120-123.
- [4] Englehart K, Hudgins B, Parker PA, et al. Classification of the myoelectric signal using time-frequency based representations. Med Eng Phys, 1999, 21: 431-438.
- [5] Chu JU, Moon I, Mun MS. A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand. IEEE Trans Biomed Eng, 2006, 53: 2232-2239.
- [6] 余龙, 吴禄慎, 张志勇, 等. 基于虚拟现实技术的多功能肌电假肢控制系统开发平台. 计算机测量与控制, 2011, 19: 609-612.
- [7] 蔡铁, 朱杰. 自动语音识别系统中的 OOV 快速拒识算法. 计算机工程, 2005, 31: 23-24.
- [8] Tatsuya H, Makoto O, Shota S, et al. Silent-speech enhancement using body-conducted vocal-tract resonance signals. Speech Commun, 2010, 52: 310-313.

(修回日期:2012-09-03)

(本文编辑:汪 玲)

· 短篇论著 ·

吸氧联合呼吸肌训练治疗慢性阻塞性肺疾病的疗效观察

李晓艳

慢性阻塞性肺疾病(chronic obstructive pulmonary disease, COPD)发病与大气污染、职业性粉尘、化学物质、病毒或细菌感染以及长期吸烟等因素有关,患者因气道阻塞、肺及胸廓顺应性下降、肺过度膨胀,使膈肌受压、低平,膈肌活动度变小、收缩力下降,呼气时气道塌陷闭塞,肺泡气体不能有效排出,从而影响机体正常气体交换^[1]。我科采用吸氧及呼吸肌训练联合治疗 28 例中、轻度 COPD 患者,发现临床疗效显著。

一、对象与方法

共选取在河南省第二慈善医院和河南理工大学附属医院治疗的男性 COPD 患者 28 例,年龄 45~60 岁,平均(55.6±4.8)岁;病程(5.5±2.5)年,其诊断均符合中华医学会呼吸病学分会 2002 年颁布的《慢性阻塞性肺病诊治指南》中关于 COPD 的判定标准^[2],病情属于轻度或中度。

入选患者均给予吸氧及呼吸肌训练,首先通过鼻导管进行低流量持续吸氧,吸氧流量为 1~2 L/min,氧浓度为 20%~29%,持续 1 h 后进行呼吸肌训练,包括:①深长呼吸训练,嘱患

者取站立位并屏气,然后进行最大幅度吸气及最大幅度呼气动作,吸气及呼气动作各持续 4~6 s,持续训练 5 min;②腹肌肌力训练,嘱患者取仰卧位,腹部放置沙袋进行挺腹练习(吸气时腹部隆起,呼气时腹部下陷),开始训练时沙袋质量为 2.5 kg,以后逐步增加沙袋重量至 5 kg 以上,腹肌练习每次持续 5 min。待上述深长呼吸训练及腹肌训练结束后再进行 1 h 吸氧治疗,吸氧参数同上。上述吸氧及呼吸肌训练每天分早、中、晚各进行 1 次。

于治疗前、治疗 6 周后进行肺功能检测,采用美国产 SENSOR MEDICS-2200 型肺功能仪及全自动血气分析仪,主要肺功能检测指标包括:用力呼气时所呼出的最大气量(forced vital capacity, FVC)、第 1 秒用力呼气容积(forced expiratory volume in one second, FEV1)、卧位动脉血氧分压(PaO₂)及二氧化碳分压(PaCO₂);其中 FEV1/FVC(%)≤70 属于 COPD 患者评定等级范围,PaO₂ 和 PaCO₂ 分别提示 COPD 患者血液中氧含量及二氧化碳滞留状况,主要反映患者肺气体交换及代谢能力^[3]。

本研所得数据以($\bar{x} \pm s$)表示,选用 SPSS 13.0 版统计学软件包进行数据分析,统计学方法选用 *t* 检验, *P* < 0.05 表示差异具有统计学意义。

DOI:10.3760/cma.j.issn.0254-1424.2012.010.023

基金项目:河南省教育厅自然科学研究计划项目(2008B890006)

作者单位:454000 焦作,河南理工大学体育学院人体科学实验室